# 6 UNCERTAINTY ANALYSIS

Uncertainty analysis, as it pertains to the SFWMM, refers to the establishment of confidence limits or uncertainty bands for model output variables as a function of the assumed confidence limits for the model input parameters. Model uncertainty analysis provides the modeler and/or decision-maker some information about the limitations of the model output. By being aware of the uncertainties associated with model output, misuse of performance measures derived from such output can be reduced. Also, similar to sensitivity analysis, future efforts on data collection can be enhanced by insight provided by the results from an uncertainty analysis. The reliability of the model is limited by the uncertainties in the output. This section summarizes some of the methods used and findings of a certainty/uncertainty analysis of the SFWMM. This chapter is entirely based on Trimble (1995a).

## 6.1 METHODOLOGY

Model output uncertainty is a function of three other uncertainties: model input data uncertainty, model algorithm uncertainty, and model input parameter uncertainty. The first two sources of uncertainty can be quantified and are partially discussed in Abtew, et al. (1993), Kone (1992), and Lal (1995). This section focuses on the last source of model output uncertainty: model input parameter uncertainty, or simply, parameter uncertainty. The approach will be two-fold:

**1.** Model output uncertainty will be estimated as a function of parameter uncertainty using a first-order approximation; and
**2.** Total output uncertainty will be determined by applying a regression analysis technique.

The first-order approximation of uncertainty associated with model output or performance measure is derived by obtaining the variance of the first-order approximation of the multi-variate Taylor series expansion of the model output or an associated performance measure. The variance, assuming an independent set of parameters, can be expressed as:

$$VAR[M] = \sum_i \left[ \left( \frac{\partial M}{\partial P_i} \right)^2 VAR(P_i) \right] \qquad \forall \quad i = 1,\ldots n \qquad (6.1.1)$$

where:

- $M$ = a model output variable or performance measure;
- $VAR[M]$ = variance in the output variable or performance measure;
- $P_i$ = $i^{th}$ parameter;
- $VAR(P_i)$ = variance of the $i^{th}$ parameter;
- $\partial M/\partial P_i$ = the sensitivity coefficient; or partial derivative of M at the mean value of input parameter i; and
- $n$ = number of parameters.

Eq. (6.1.1) states that output uncertainty due to parameter uncertainty can be estimated by multiplying the variance of each parameter by the square of the parameter's sensitivity coefficient and summing the values for all parameters. This method allows for estimation of percentage contribution of each parameter variance to the total model output variance.

Calculations associated with sensitivity coefficients are presented in the previous chapter. By assuming each model input parameter as a normally distributed random variable, the variance of parameter i can be approximated by the following equation (Loucks and Stedinger, 1994):

$$VAR(P_i) = [ (P_{95} - P_{05}) / 3.3]^2 \qquad (6.1.2)$$

where:

$P_{95}$ = upper limit of the 90% confidence interval; and
$P_{05}$ = lower limit of the 90% confidence interval.

Therefore, when a normal distribution represents the likelihood of each parameter value, the distance between the parameter mean value and the upper or lower limit of the 90 percent confidence band is 1.645 standard deviations.

In order to express total model output uncertainty, i.e., the combined effects of the three sources of uncertainties mentioned earlier, Flavelle (1992) suggested a methodology that uses linear regression with observed data as independent variable and simulated data as dependent variable. Regression coefficients are used to estimate model uncertainty. Although regression analysis does not allow the sources of error to be directly calculated, some inferences can be drawn about the relative uncertainties (parameter uncertainty versus data/algorithm uncertainty) by comparing the results of this method from those obtained using first-order approximation. Also, the latter method assumes that rainfall, evapotranspiration and historical structure flow information are reliable data with insignificant, if any, errors. Regression analysis incorporates all types of errors.

Uncertainty bands for an individual predicted value can be estimated using the following relationship (Lal, 1994; Walpole and Myers, 1990):

$$Y = a + bx \mp s_e [ 1 + 1/n + (x - x_m)^2/\sigma^2]^{.5} t_{.95} \qquad (6.1.3)$$

where:

$Y$ = model simulation value to be predicted using the regression equation;
$a$ = regression coefficient representing the Y-intercept of the regression line;
$b$ = regression coefficient representing the slope of the regression line;
$x$ = observed value;
$x_m$ = mean of the observed value;
$\sigma$ = variance of x;
$t_{.95}$ = student-t value for a one-tail 95 percent confidence interval; and
$s_e$ = square root of the unbiased estimate of the standard error of the regression equation.

The estimator $s_e^2$ is equal to the sum of the squares of the errors about the regression line divided by the number of degrees of freedom which, in turn, is equal to the number of points used in the regression minus two.